MODIS Aerosol Optical Depth Bias Adjustment Using Machine Learning Algorithms

Arif Albayrak[1,2], Jennifer Wei[1,2], Maksym Petrenko[3], and Gregory Leptoukh[4]

1,2 - NASA GES DISC, ADNET , 3 – NASA, UMCP, 4 – NASA GES DISC

Poster: MODIS Science Team Meeting 2012

Goal and Application

Goal : The goal of this work is to investigate MODIS dark target aerosol products (MOD04, MYD04, v051) compared to AERONET quality assured L2 aerosol, and then to adjust the systematic biases and errors using Machine Learning algorithms (i.e., Neural Network).

Application : The results of bias adjustment for MODIS Terra and Aqua are incorporated into the Giovanni AeroStat as part of the NASA ACCESS project. (http://giovanni.gsfc.nasa.gov/aerostat/)

Background

Over the past decade, global aerosol observations have been conducted by space-borne sensors, airborne instruments, and ground-base network measurements. Unfortunately, quite often we encounter the differences of aerosol measurements by different well-calibrated instruments, even with a careful collocation in time and space. The differences might be rather substantial, and need to be better understood and accounted for when merging data from many sensors. The possible causes for these differences come from instrumental bias, different satellite viewing geometries, calibration issues, dynamically changing atmospheric and the surface conditions, and other "regressors", resulting in random and systematic errors in the final aerosol products.

Methodology

Neural Networks: Neural Networks (NN) is applied to remove biases and the systematic errors from MODIS (both Terra and Aqua) aerosol product, using the Aerosol Robotic Network of sun-photometers (AERONET) as a baseline for evaluating the MODIS aerosol products globally. The overall NN algorithm is described as following:

- Network Architecture: Feed-Forward (no cycles)
- Optimization Algorithm: Back Propagation
- Hidden Layers: One
- Nodes: 40
- Number of Regressors
  - Dark Target Land: 15
  - Dark Target Ocean: 14

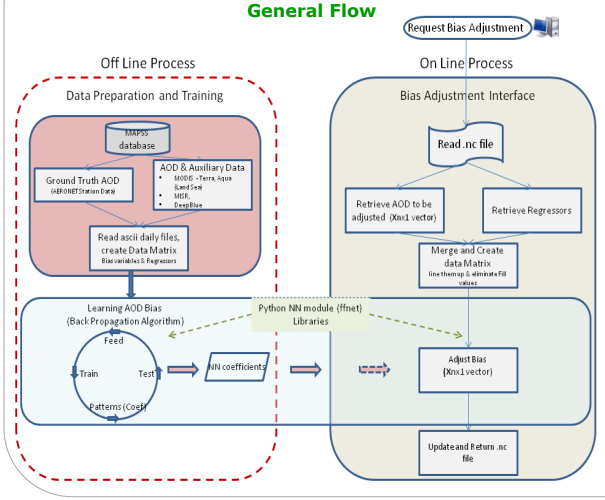
Code: Python and ffnet module by Merek Wojciechowski (Fortran + f2py)

Statistical Methods and Visualization Tools:
Because of high numbers of data with different regimes, scatter plot with 1:1 line is not enough to draw conclusions. Results are further supported with

- Standard and Non-standard tests for analysis
- Correlation coefficient analysis
- Box plots

Neural Network Model and Results

General Flow



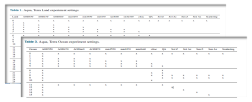
Data Set and Regressors For Training

Training and testing data set for Terra-Aqua MODIS (ocean & land), is prepared by using Multisensor Aerosol Products Sampling System (MAPSS) database over 10 years.

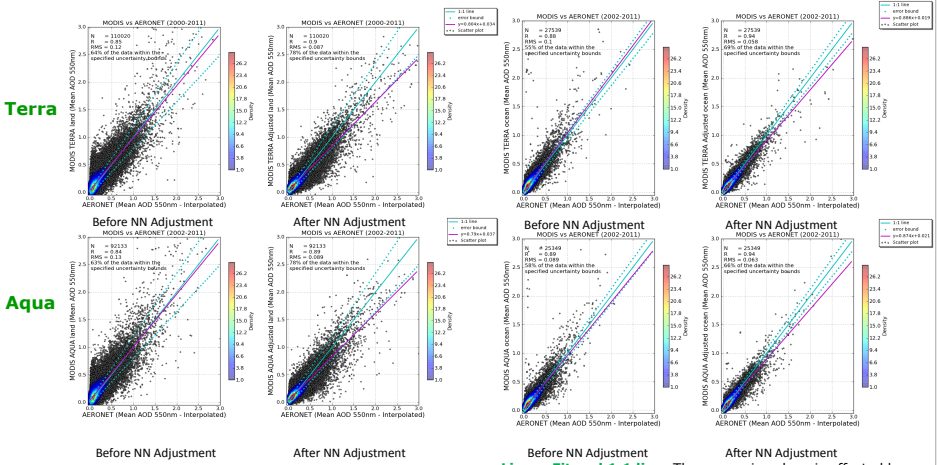
	Num of Data Total	Number of Data After filtering	Number of Trained Data	Number of Tested Data
Aqua MODIS Ocean	27434	27349	24615	2734
Aqua MODIS Land	100465	96196	86577	9619
Terra MODIS Ocean	30716	30622	27560	3062
Terra MODIS Land	123946	118926	107034	11892

MODIS Common Regressors:
Target: mean AOD0550intrp (for training only)
Geometry: SolarZenith, SolarAzimuth, SensorZenith, SensorAzimuth, ScatteringAngle
Ocean: mean AOD0550, mean AOD0470, mean AOD0660, meanAOD0870, mean mref0550, meanmref0470, mean mref0660, mean cfrac, QAavg
Land: mean AOD0550, mean AOD0470, mean AOD0660, mean mref0470, mean mref0550, meansurfre0660, mean surfre0470, mean surfre2100, mean cfrac, QA

Preliminary Tests for Regressors



Bias Adjustment Results For AQUA and TERRA (Global)



Error Cones are the confidence intervals defined below:

- Land = 0.15* Γ +/- 0.05
- Ocean = 0.05* Γ +/- 0.03
- The more the number of data in the cone, the better the results

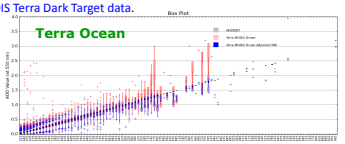
Linear Fit and 1:1 line: The regression slope is affected by high AOD points while the majority of measurements are with low AOD. The NN approach treats both equally, so when the majority of the low AOD points are adjusted for their negative bias, the overall slope tilts down away from the 1:1.

Data Analysis

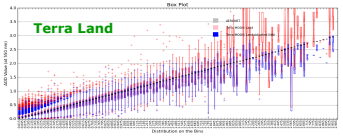
Before and after NN adjustment

Notched Box Plot (NBP) is drawn separately for AERONET and MODIS data for the each bin. Each bin represents both AERONET and corresponding MODIS aerosol data with a 0.03 interval. Red boxes show original collocated MODIS Terra Dark Target data, and blue boxes are bias adjusted collocated MODIS Terra Dark Target data.

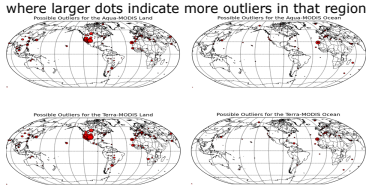
a)



b)



Outliers geolocations from AERONET stations
where larger dots indicate more outliers in that region



- Both correlation coefficient and standard deviation are improved after NN application.
- Independent analysis showed that there are 2 possible regimes in the data set causing the slope to move away from the 1:1 line. As a result, multiple statistics need to be used for realistic evaluation.

Summary of the Improvements

	RMS	Correlation Coefficient	increase of data in the cone (Percent)
Aqua Land	0.13 to 0.091	0.86 to 0.91	61 to 79
Aqua Ocean	0.09 to 0.056	0.87 to 0.94	58 to 70
Terra Land	0.13 to 0.088	0.87 to 0.92	62 to 78
Terra Ocean	0.096 to 0.053	0.88 to 0.95	55 to 75

Current and Future Work

Currently similar work has been done for MODIS Deep-Blue and MISR AOD. In the future we plan to:

- Consider multi regimes in the data sets: Run the same NN system on the global clusters that are obtained from multi-dimensional data sets.
- Combine regressors that are obtained from MISR and MODIS.
- Adjust biases for multiple data sets from different sources to a single reference and then merge them to extend coverage.

Acknowledgement

Support for the development of this on-line platform for statistical intercomparison of aerosols has been provided by NASA HQ (PM: Stephen Berrick) through the ROSES 2009 ACCESS Program (PI:Gregory Leptoukh). The AERONET data is a contribution by the International AERONET Federation (PI: Brent Holben).